

University of Colorado, Boulder CU Scholar

Undergraduate Honors Theses

Honors Program

Fall 2014

User-Adjusted Settings for Music Listening with a Simulated Hearing Aid App: Effects of Dynamic Range Compression, Data-rate and Genre

Caleb Kronen

caleb.kronen@Colorado.EDU

Follow this and additional works at: http://scholar.colorado.edu/honr_theses



Part of the [Speech and Hearing Science Commons](#), and the [Speech Pathology and Audiology Commons](#)

Recommended Citation

Kronen, Caleb, "User-Adjusted Settings for Music Listening with a Simulated Hearing Aid App: Effects of Dynamic Range Compression, Data-rate and Genre" (2014). *Undergraduate Honors Theses*. Paper 747.

This Thesis is brought to you for free and open access by Honors Program at CU Scholar. It has been accepted for inclusion in Undergraduate Honors Theses by an authorized administrator of CU Scholar. For more information, please contact cuscholaradmin@colorado.edu.



User-Adjusted Settings for Music Listening with a Simulated Hearing Aid App: Effects of Dynamic Range Compression, Data-rate and Genre

Caleb J.N. Kronen

Department of Speech Language and Hearing Sciences,
University of Colorado at Boulder

November 2014

Thesis Advisor

Dr. Kathryn Arehart, Department of Speech Language and Hearing Sciences

Committee Members

Dr. Kathryn Arehart, Department of Speech Language and Hearing Sciences
Dr. Tammy Fredrickson, Department of Speech Language and Hearing Sciences
Dr. Daniel Jones, Honors Program

Introduction

Listening to music through hearing aids or through an application (“app”) designed for a personal device (i.e. iPad, iPhone, etc.), as shown in Figure 1, involves several factors. A musical listening experience is affected by a) the nature of the musical signal b) the manner in which the app or hearing aid processes the signal c) the specific hearing impairment of the listener and d) the musical and/or audio engineering background of the listener.

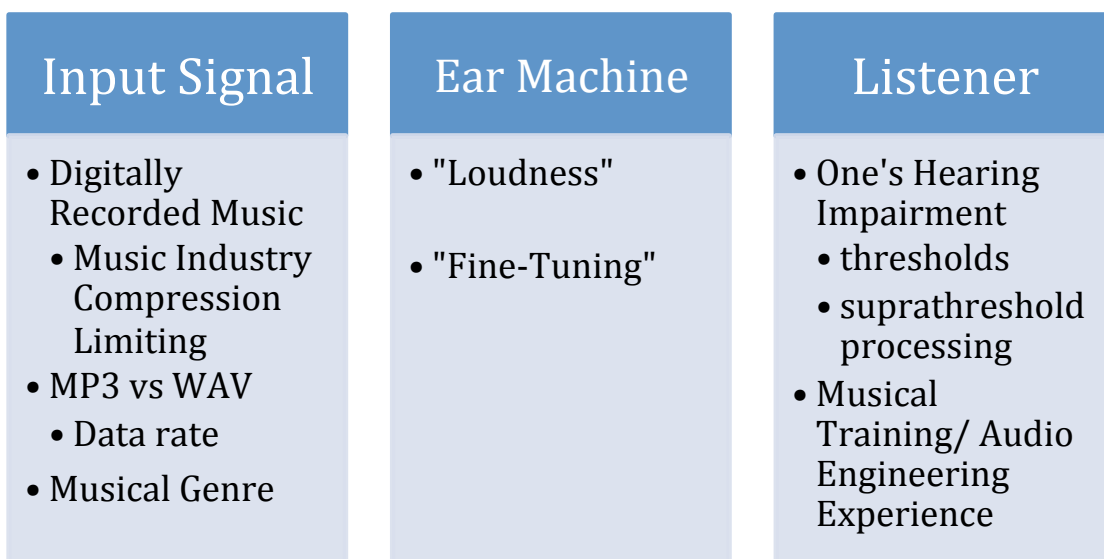


Figure 1. The three-step process of listening to music through hearing aids or simulated hearing aid apps.

There are many people who use hearing aids, yet 43% of hearing-aid users do not find that their hearing aids make a difference in their music enjoyment (Leek et al. 2008). Unfortunately for individuals with hearing impairment who enjoy listening to music, hearing aids have been largely aimed to

process speech. The significant acoustic dissimilarities between music and speech must be considered by current hearing aid manufactures in order for digital signal processing to increase the overall sound quality of music. Analyses show that 30-40% of hearing aid users in the United States most likely own a smartphone (personal listening device) and that will likely double within 5 years (Sabin, et al 2013). Based on consumers' lack of enjoyment with listening to music through hearing aids and the increasing number of individuals who use personal listening devices (e.g. iPad, iPhone), researchers have been developing applications for smartphones and tablets that are designed to increase the enjoyment of listening to recorded music over hearing aids or headphones.

Pathways of the musical signal

Individuals with hearing impairment can listen to music via different pathways. Musical input signals can originate from live music or digitally recorded music which can be delivered through loud speakers and/or through direct audio input (e.g. an "app" and an iPad). Both live and digital music are affected by factors in a listening environment or by sound transmission pathways in the listening environment. Distortions of the signal may occur at different stages on the pathway from the sound source to the listener's ear, hearing aid, or app.

Listening to Digitally Recorded Music

According to a recent NPR story, the Consumer Electronics Association reported that 90% of consumers say sound quality is the most important part of the recorded music listening experience (Markwalter, 2011). Because of this

consumer preference, maximum sound quality has become a goal of many recording artists and audio engineers.

This study focuses on the quality of music processed through a recently developed app called Ear Machine. Ear Machine is an application for the Apple® iPhone and iPad that was created with funding from the National Institutes of Health and the Capita Foundation. Ear Machine allows people with hearing impairment to adjust processing settings while listening to recorded music over in-ear headphones on their iPhone, iPad or iPod Touch. The development of a hearing-aid-app is timely because of the large number of hearing aid users estimated to own smart phones.

Audio signals, such as MP3 or WAV files, are sent from a source (e.g. smart phone or tablet) through signal processing that is applied by simulated hearing aid applications. Because of this direct route of sound, environmental distortions are minimized while using Ear Machine. Regardless of the route, before digitally recorded music is sent to a loud speaker or Ear Machine, the signal is encoded and also undergoes other processing during the recording process. This encoding process and other signal processing during the recording process can introduce their own forms of distortion which may or may not decrease overall perceived sound quality.

Literature Review

Input Signal

Acoustics of the music signal:

The dynamic range of a musical piece refers to the decibel difference between the excerpt's greatest level and lowest level. Music is more intense than speech and is also more dynamic (Chasin, 2003; 2006; 2010). Crest factor is described by the ratio of peak level to the average intensity level. The value of the crest factor represents how intense each peak is in the given sample of sound. By indicating how extreme the peaks are in a waveform, crest factors can express how dynamic a segment of music is. The difference in dynamic range between rock and classical music varies across the frequency spectrum. Specifically, dynamic range difference plots reveal that rock music has much more energy at low frequencies (under 200 Hz) and high frequencies (over 8,000 Hz).

Not only does music contain a wide range of intensity levels but it also contains a wide range of spectral content. From one musical instrument to another, there is more variability in the range of spectral content than from one human voice to another human voice. Thus, music contains a very wide range of frequencies that may be simultaneously present due to the presence of a range of instruments that vary in size and physical properties. The range of frequencies present in music is also dependent on the musical genre. As shown in Figure 2, rock music has more high frequency emphasis due to rock's heavy use of crash, ride and hi-hat cymbals (Bregitzer, 2009).

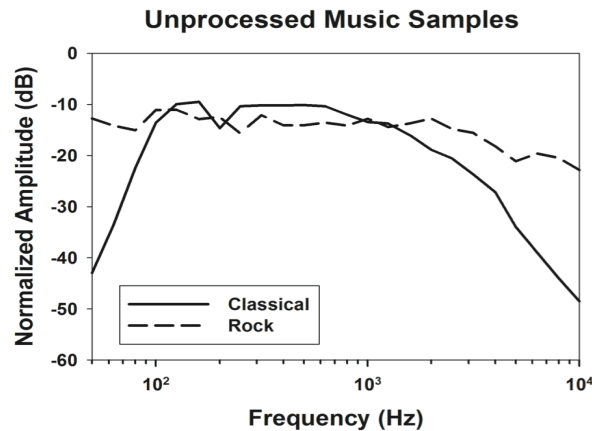


Figure 2. The long-term spectra of the unprocessed rock and classical samples.

Dynamic range at low and high frequencies varies for genre. For example, Croghan et al (2014) documented that rock music has much wider dynamic range at low and high frequencies than the classical music. However, classical music commonly contains very high intensity and low intensity segments within a piece of music, whereas rock music typically does not have such a dynamic range of intensities overall. Because of this difference, classical music usually has a crest factor of 10.25 dB while rock music has a crest factor of 9.15 dB.

Dynamic-Range Compression (DRC) used in the music industry

In order to increase loudness, music industry dynamic range compression (DRC) is designed to decrease the difference between the greatest level and lowest level sounds in an excerpt. By decreasing the dynamic range of a stimulus, high level sounds are reduced in amplitude and lower level sounds are increased in amplitude. Unlike live music, digitally recorded music undergoes various signal processing stages throughout the recording process. For example, there are multiple instances during the music-industry recording process where DRC is applied (Bregitzer, 2009), mainly because DRC is implemented as a

creative tool by controlling the attack times and levels of various instruments

(Bregitzer, 2009). Based on information from Katz, & Katz (2003), it is clear that DRC is habitually added to vocals, drums, guitar and most other stem tracks.

Typically, the input signal of live music (over sound field) is uncompressed and minimally altered; however, because the dynamic range for live music is usually 100dB which is much greater than the dynamic range for digitally recorded music (Chasin, 2006), DRC is rarely added to live signals but is commonly applied to recorded music. Creating one hearing aid setting for music is difficult because depending on the genre and instrumentation of a piece, audio engineers apply large amounts of DRC to one song, while applying minimal amounts of DRC to another song.

Compression in Music in Hearing Aids

Hearing aids are complex devices that implement several layers of signal processing including amplification, compression, frequency lowering, and noise reduction (Kates, 2008). This paper focuses on the compression ratios of wide-dynamic range compression (WDRC).

Compression is commonly implemented in hearing aids during music listening and speech listening even though the two stimuli require different amounts of compression. For example, both Arehart et al (2011) and Hansen (2002) reported that higher compression ratios resulted in decreased music quality compared to lower compression ratios. Because music is usually more intense than speech, music signals are commonly distorted in hearing aids (Chasin 2003, 2006, 2010).

As expressed earlier in this paper, recorded music often undergoes DRC in the music industry and then hearing aids also apply WDRC. In order to examine the roles of input-signal properties and hearing-aid processing in the perception of recorded music, with an emphasis on the effects of dynamic-range compression, Croghan et al (2014) collected perceptual quality ratings. Results showed that listeners perceived maximum quality of recorded rock and classical music when WDRC was linear (a compression ratio of 1:1). The findings from Croghan et al (2014) explain how WDRC applied to music by hearing aids should be linear in order to accommodate for the effects of DRC that is applied in the music industry.

Using 18 simulated hearing aid conditions, Croghan et al (2012) found that due to acoustic content differences and recording process differences, quality ratings of rock and classical music were affected differently by hearing aid WDRC. Specifically, classical music needed less hearing aid WDRC than rock music for the highest quality ratings. By varying the amounts compression, Croghan et al (2012) found that music with less hearing aid WDRC received the highest music quality ratings.

When the signal is amplified, WDRC applies more gain to low-level sounds and less gain to high-level sounds. This processing increases audibility and also eliminates peak clipping. Kates (2010) explained that WDRC algorithms create a relationship between audibility and distortion such that decreasing distortion usually decreases audibility. The distortion that occurs from compression degrades the fidelity of the temporal envelope (Jenstad & Souza

2005). Distortions from compression also have shown to flatten the frequency spectrum (Plomp 1988).

Higgins et al (2012) examined listener preferences between music with WDRC applied versus music with adaptive dynamic-range optimization applied. Quality ratings of rock, classical and jazz music were higher when adaptive dynamic-range optimization was applied versus WDRC. Another recent study Davies-Venn et al, (2007) concluded that listeners with hearing loss hear better music sound quality from WDRC as opposed to peak clipping and compression limiting; however, the variance in the WDRC configurations between Davies-Venn et al (2007) and Higgins et al (2012) left the field unclear concerning the benefit that dynamic-range compression can provide for increased music quality. The conclusions of Davies-Venn et al are consistent with the findings of Hawkins and Naidoo (1993). Hawkins and Naidoo (1993) reported that subjects with hearing impairment favored music that was compressed versus music that had peak clipping.

Although WDRC that is implemented in hearing aids is intended to overcome the issues associated with reduced dynamic range and abnormal loudness growth seen with cochlear hearing loss (Moore and Glasberg, 1997), WDRC also warps the signal envelope, reducing the spectral and temporal contrasts that listeners commonly utilize in music perception (Plomp, 1988). While WDRC found in hearing aids is thought to increase sound quality by raising musical signals above threshold (Villchur, 1973), a high amount of compression may diminish the perceived quality of music.

Hansen et al (2002) found that listeners with hearing loss perceived music signals differently than individuals with normal hearing. Souza et al (2007) showed that some types of hearing aid processing affects the perceived pleasantness of music samples more than other types of hearing aid processing. Tan et al (2004, 2008) compared quality ratings of stimuli of different amounts of peak clipping and instantaneous compression. Both studies found that increased amounts of clipping and increased amounts of WDRC result in lower quality ratings for both listeners with normal hearing and listeners with hearing impairment. Although normal-hearing and hearing-impaired listeners provide higher music quality ratings as amount of compression was lowered (Van Buuren et al, 1999), there has not been a study that examines this relationship during the use of an app that simulates a hearing aid.

Data Rate, Encoding, and Latency

The number of times at which the amplitude of an analog signal is converted into data (comprised of 0's and 1's) per second is what defines the sampling rate. Due to the effects of decreased outer and inner hair cell function in listeners with hearing loss, the effect of a slower data rate (e.g. 22,050 Hz) on their ability to perceive sound quality is both significant and difficult to predict (Rix, 2006). It is difficult to predict how distortions, that are present in lower data-rate digital music, change the listening experience for listeners with hearing loss.

Latency can be thought of as how long it takes a sound to move from its source, become processed by the app, and then be delivered to one's ears. While listening to high-resolution music, latency can be thought of as how long it

takes for the device to send the data through a processor and then through earphones until it reaches the listener's ear. According to one of the creators of Ear Machine, "it is generally thought that latency below about 30 ms is acceptable for a hearing app," (Sabin, 2013). It takes more time to process higher resolution audio files because there is more information to process per second and Ear Machine audio quality can be compromised by distortions caused by a lack of processing speed.

Pras et al (2009) showed that musical genre played a role in the perception of hearing quality differences. Specifically, electric music (versus acoustic music) provided more salient quality changes to the listeners when sampling rate was changed. Similarly to Pras et al (2009), Ruzanski (2006) found that musical genre has a significant effect on the quality changes that occur from different data rates (kbps).

Musical recording files are commonly saved in mass quantities on smartphones and portable listening devices. Since the introduction of the MP3 format in 1991, MP3's reputation has slowly converted from a low resolution storage file format to the most common commercial music file type for portable music devices and smart phones. As MP3's increase in popularity and physical CD's containing WAV files decrease in popularity, developers of app's such as Ear Machine must consider how MP3's can introduce new distortions that are audible to users.

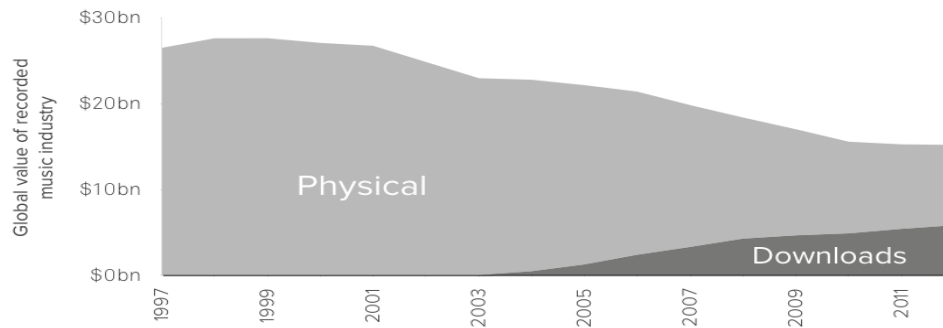


Figure 3. Global value of recorded music industry by year for physical CD sales and download sales. (Source: www.CDbaby.com)

The sampling rate of the input signal also changes the way in which a Ear Machine processes and transfers a musical signal. Specifically, the signal encoded in an MP3 is already distorted due to the detrimental encoding process and commonly low sampling rate of 22,050 Hz. Because of MP3 files' pre-distorted signal, the files' low sampling rate can have a deleterious effects on quality when listening through Ear Machine.

While the input signal is transferred through a digital-to-analog-converter (DAC) and then sent to Ear Machine, the signal can be distorted due to the quality of the sound card inside the sound source (e.g. smart phone, computer, iPad). Depending on the quality of the sound card within the device, the signal fidelity is altered (Bregitzer, 2009).

CD-quality audio (WAV files) contain an un-altered version of audio information. In contrast to CD-quality audio (WAV files), MP3 files are created in a process that removes data from an audio recording that does not play a purpose in the audio (e.g. noise or information expressing an absence of sound) (Rix, 2006). MP3 encoders are designed to process sound information based on

a psychoacoustic model which takes in to account the presence of frequency and temporal masking. During this process, distortions to signal fidelity occur, including changes to the fine structure and envelope of the signal (Rix, 2006). The higher the kbps of the MP3, the less distortion that will occur in the audio file from the rounding of values that is performed by the algorithm.

The audio information that is encoded in an MP3 file is typically an 8-bit-22,050 Hz (sampling rate); however, less commonly, MP3 files can be a 16-bit 22,050 Hz (sampling rate). When MP3 encoding occurs, the lower the sampling rate, the more information is removed from the original audio signal. In order to increase the amount of storable music on a smartphone or iPad, MP3 files reduce the kilobits per second (kbps) and as a result, MP3 files with 44,100Hz sampling rates still have less kbps than WAV files with a 44,100Hz sampling rate.

MP3 encoders take psychoacoustics into account in order to extract information in such a way that listeners will not hear audible distortions. Audible or not, more distortions and degradation of information occur when MP3s are encoded compared to CD-quality audio.

In the majority of MP3 files, high-frequency content is lost due to a common sampling rate of 22,050Hz. The Nyquist–Shannon sampling theorem claims that the sampling frequency must be more than twice the maximum frequency in the desired file.

Kuk (2010) states that while a higher sampling rate (e.g. 44,100Hz) usually corresponds to a clearer signal with better sound quality, unfortunately, more bits also mean more calculations, more memory, and longer time to

transmit. More information to transfer can cause a perception of diminished quality due to latency issues in the play-out of the file.

Pras et al (2009) looked at individual listeners' abilities to hear quality differences between CD and MP3 quality. They found that listeners can hear a difference between CD quality and MP3 quality between 96-192 kbps, but only expert listeners with formal audio engineering experience can hear differences between CD quality and MP3 quality between 256-320 kbps.

	Expert Listener	Novice Listener
96-192 kbps	Yes	Yes
256-320 kbps	Yes	No

Table 1: Data from Salimpoor (2006)

Expert listeners can hear differences between CD quality and MP3 quality between 256-320 kbps but novice listeners cannot.

As shown in Table 1, Salimpoor (2006) reported that listeners without audio engineering experience were only able to hear a significant change from CD quality to low bit rate MP3's. Sutherland et al (2007) discovered that experienced audio engineers preferred CD quality over MP3s even when the MP3 file has a high kbps (320kbps).

Gan & Kuo (2007) report that the rate of 128 kbps is the most commonly used. After taking a survey of the kbps data rate of MP3 downloads from three MP3 distributors (CD

<i>iTunes</i>	<i>Amazon</i>	<i>CD Baby</i>
128 kbps	128 kbps	128 kbps

Table 2: The average data-rates for MP3 files sold by iTunes, Amazon, and CD Baby

Baby, iTunes, and Amazon), the averages in Table 2 were obtained. Based on these averages, a listener without audio engineering experience would not be able to decipher the quality difference between CD quality and MP3 quality of an iTunes MP3 or an Amazon MP3 (Pras et al, 2009). Both Advanced Audio Coding and MP3 distort the frequency spectrum of music significantly, even at a data rate of 320kbps (Atkinson, 2008) which is much higher than the data rate of the majority of MP3 files. Much of this distortion comes from the sampling rate of 22,050Hz.

Audiologist-Driven Versus Patient-Driven Fine Tuning of Hearing Instruments

In the current model of hearing aid fitting, an audiologist takes a patients audiogram and uses manufacturer fitting software and audiologist adjustments to reach maximum sound quality for the user. This process may take several visits. A potential alternative model is to reassign adjustments to the listener, which may now be more feasible as smartphones and other personal listening devices have gained popularity.

Boymans & Dreschler (2011) examined listeners' speech perception in quiet, in noise, and in time-reversed speech with audiologist-driven versus patient-driven fine tuning of the same hearing aid. Participants performed better with the audiologist-driven settings for speech perception in quiet and in time-reversed speech. While perceiving loud sounds, listeners rated the overall quality of higher when fine tuning was patient-driven. Overall preference for 67% of the participants was with audiologist-driven settings. Particularly for 1000 and 2000 Hz, the audiologist-driven approach generated a higher gain than the patient-driven approach.

Processing adjustments using the Ear Machine App



By adjusting the loudness and fine-tuning controllers in the app, a patient can be adjusting over 100 different values that control digital signal processing parameters (gain, compression ratio, and maximum power output). Any change in either of the controllers creates a new and unique listening settings. When a user moves the 'loudness'



controller upwards, the listener is applying a more energetic prescription (e.g. a nonlinear gain prescription that initializes at 1/2-gain, instead of 1/3-gain). The loudness controller adjusts WDRC compression ratio, gain and maximum power output. Ear Machine uses WDRC with a fast attack and a slow release time.

When users adjust the fine-tuning controller, they are applying a nonlinear prescription with either more or less high-frequency emphasis, using 1000 Hz as the pivotal point. The fine-tuning controller controls the gain that is applied in 12 frequency bands. Whenever either of the controllers is adjusted, the app will create a new array of gain values, within its 9 WDRC bands for low-, mid- and high-level sounds. When the fine-tuning controller is moved upwards,

more gain is applied to frequencies above 1000 Hz, and when the fine tuning controller is moved downwards, more gain is applied to frequencies below 1000 Hz.

Amlani et al (2013) used speech-in-noise tests and surveys to study how individuals use Ear Machine. Amlani et al (2013) explains that in a noisy environment, subjects found the smartphone application provided better sound quality than a Unitron Shine and Moda II 312 behind-the-ear hearing aids. By revealing that the app has comparable electroacoustic characteristics to conventional hearing aids, Amlani (2014) explained why subjects observed somewhat analogous functioning between the app and the hearing aids. Even though the iPhone, iPad, and iPod Touch processors are so powerful, due to the minimal amount of amplification Ear Machine offers, traditional hearing aids are still the best option for individuals with severe hearing impairment. Another issue for some users is that Ear Machine only works with hardwired, in-ear headphones.

In phase one of Ear Machine's research, subjects listened to pre-recorded speech-in-noise and music passages. Using an iPod Touch, 41 subjects with hearing impairment from mild to severe manipulated the Ear Machine controllers, simulating a 9-channel WDRC hearing aid, in order to adjust for maximum intelligibility and perceived sound quality. The results of the study revealed that participants are able to make reliable Ear Machine adjustments in about 10 seconds. Phase one also revealed that even if subjects do not begin with settings suitable for their specific degree and configuration of loss, the subjects set gain

with about -5 dB re: The National Acoustic Laboratories' (NAL) prescribed gain, which is close to how the subjects of Keidser et al. (2008) preferred gain about -5 dB re: NAL. For rock, classical and jazz music, the subjects preferred self-adjusted settings better than the settings prescribed by NAL. Subjects also preferred user-adjusted music settings more than audiologist-adjusted settings for music. Phase one of Ear Machine research shows that users adjust controllers reliably and that they found the controllers to be helpful and useful (Sabin et. al, 2013)

Input file DRC has not been specifically applied to music listening apps. However, depending on the amount of DRC that has been applied to an audio file during the music recording process, the WDRC that is applied by an app or a hearing aid can create two-fold effects of compression. If a rock song has heavy DRC applied during the recording process and a classical song has barely any DRC applied to it during the recording process, it might be useful for a patient to be able to adjust the WDRC parameters in an app differently for different songs.

Purpose

With the number of hearing aid users who own smart phones on the rise, the heavy use of compression limiting in the music industry, the rise in use of MP3 files, and the findings from phase one of Ear Machine's research, it is important that the following questions are answered in order to determine the validity of user adjustments on Ear Machine, and to understand what factors of the input signal affect these adjustments:

- Do listeners with normal hearing and listeners with hearing impairment adjust controllers reliably (within day and across day)?

- Do listeners with normal hearing and listeners with hearing impairment adjust controllers differently for unprocessed, mildly compressed and heavily compressed music files?
- Do listeners with normal hearing and listeners with hearing impairment adjust controllers differently for MP3 and CD quality files?
- Do listeners with normal hearing and listeners with hearing impairment adjust controllers differently for rock and classical music?

With the answers to these questions, the aid provided by apps such as Ear Machine can be validated, and patients' use of these apps can be guided by evidence.

Methods

Stimuli


Unprocessed Recordings

The music stimuli included two recordings: one classical sample and one rock sample. The classical sample was an excerpt from “Overture to the Magic Flute” by W.A. Mozart and was performed by the University Symphony Orchestra at the University of Colorado-Boulder. The rock sample was an excerpt from “Anything At All” by Mere. The recorded stereo files were obtained directly from the recording engineers in their final mixes prior to the mastering stage, with no DRC applied. The original unprocessed signals were sampled at 24 bit, 44.1 kHz. Samples of approximately 13 seconds in duration were selected from the recordings, at a point consistent with musical phrasing (see Figure 2).

Using a commercial sound-editing package (Adobe Audition, Version 1.5), the stimuli were made equal in terms of overall root-mean-square (RMS) level and exported as both WAV (1411 kbps) and MP3 (128 kbps) files. These files were copied to an iPad where were be accessible by Ear Machine.

Procedure

After analyzing objective measures of the music stimuli with customized analysis programs such as amplitude histograms and crest factor measures, perceptual user-adjustments were collected. Adjustments made by users were analyzed to gain a better understanding of how preferred settings change based on genre, input compression limiting, or data rate of a given music file.

Listeners were seated in a sound attenuating booth and were presented music samples through stock Apple  headphones (see Table 3 for specifications) which were connected to an iPad set to airplane mode. Numeric verbal cues were given before each condition in order to differentiate conditions and randomize the order of each condition.

Headphone specifications:

Using Impedance: 23 OHMS
Sensitivity: 109 DB
Frequency (high): 21 kHz
Frequency (low): 5 Hz
Operating Principle: Open Air
Weight: 10 g
Cable Length: 1.395 m

Table 3: Headphone specifications for the stock Apple ear-bud headphones.

Listeners adjusted the loudness and the fine-tuning controllers in Ear Machine for maximum perceived quality for each condition. Subjects were given the following instructions: “While listening to each musical excerpt, please adjust the loudness

controller and the fine-tuning controller for best overall sound quality.” After each user adjustment, the researcher recorded the two parameters of the loudness

Condition	Genre, Compression Limiting, File type	kbps
1	Rock UNP WAV	1411
2	Rock UNP MP3	128
3	Rock ML WAV	1411
4	Rock ML MP3	128
5	Rock HL WAV	1411
6	Rock HL MP3	128
7	Classical UNP WAV	1411
8	Classical UNP MP3	128
9	Classical ML WAV	1411
10	Classical ML MP3	128
11	Classical HL WAV	1411
12	Classical HL MP3	128

Table 4: List of 12 conditions

controller and the fine-tuning controller prior to the initiation of the next condition (see Table 4 for the twelve conditions). The researcher did not assist the subject in moving the controllers. The experiment required approximately a two-hour time commitment from each

participant, spread out over 2

sessions (see Table 5 below for an outline of experimental sessions).

Session	Duration	Explanation of procedures
1	.5 hours	The subject received a hearing test from an audiologist. Subjects were presented a consent form.
1	.5 hours	Subjects listened to one practice playlist, containing all twelve conditions. During the practice playlist, participants adjusted the loudness controller and the fine-tuning controller to reach maximum perceived sound quality for each of the twelve presented musical segments. After the practice session, the subject was presented with the same twelve stimuli in a different randomized order. For each subject, the researcher recorded two test trials of adjusting Ear Machine. Subjects were asked to “please adjust the loudness controller and the fine-tuning controller for best overall sound quality”.
2	1 hour	Subjects listened to one practice playlist, containing all twelve conditions. During the practice playlist, participants adjusted the loudness controller and the fine-tuning controller to reach maximum perceived sound quality for each of the twelve presented musical segments. After the practice session, the

		subject was presented with the same twelve stimuli in a different randomized order. For each subject, the researcher recorded two test trials of adjusting Ear Machine. Subjects were asked to “please adjust the loudness controller and the fine-tuning controller for best overall sound quality”.
--	--	---

Table 5. Outline of experimental sessions

Results

Within-Day Reliability

For within-day adjustments, subjects reliably adjusted loudness ($R^2 = 0.952$). This correlation shows that users can reliably adjust the loudness controller from one adjustment to another, within a given day or session. Within-day fine-tuning user-adjustments were somewhat reliable ($R^2 = 0.726$) but less reliable than within-day loudness user-adjustments. Both Figure 5 and Figure 6 show the user-adjustments within a day (or session).

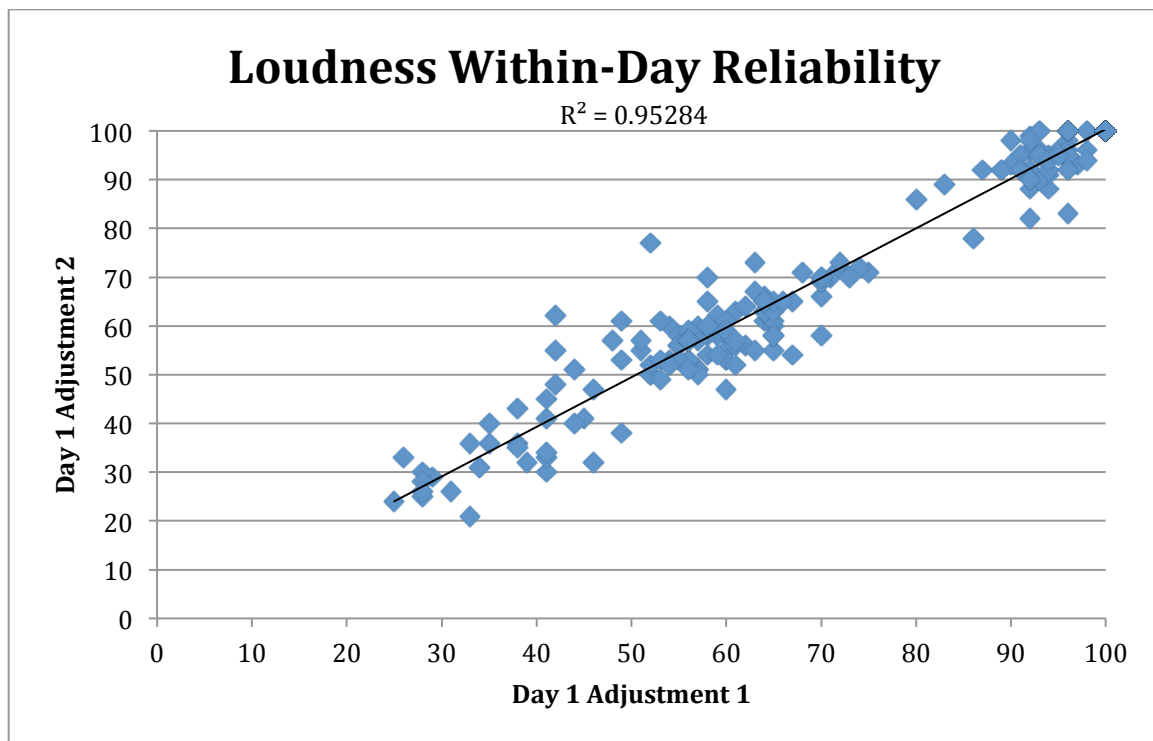


Figure 5. Scatter plot of the first adjustment on day one and the second adjustment on day one.

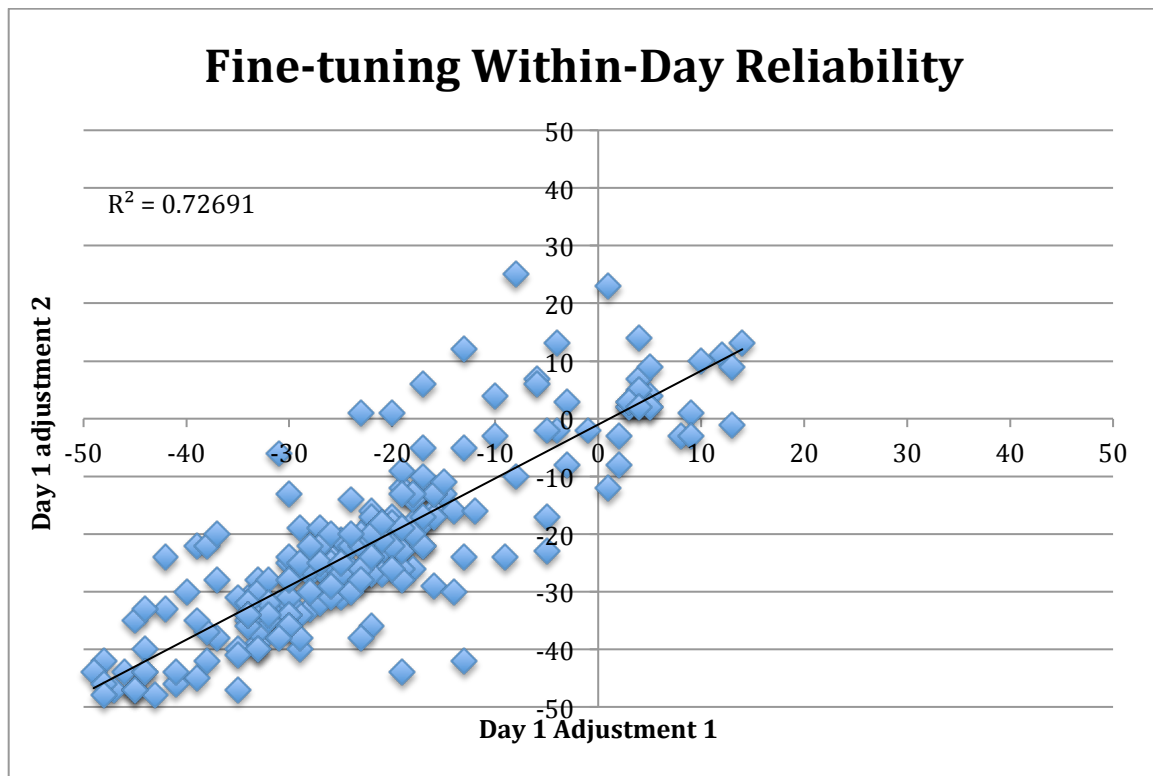


Figure 6. Scatter plot of the first adjustment on day one and the second adjustment on day one.

Across-Day Reliability

Across-day user-adjustments of the loudness controller were also reliable. Ratings on day one and day two are highly correlated ($R^2 = 0.934$), exhibiting that users of Ear Machine can reliably adjust the loudness controller between different days. Since users can adjust loudness reliably, one can speculate that individual user preferences for of Ear Machine are not randomly chosen by users and instead, users have the ability to adjust loudness in Ear Machine systematically.

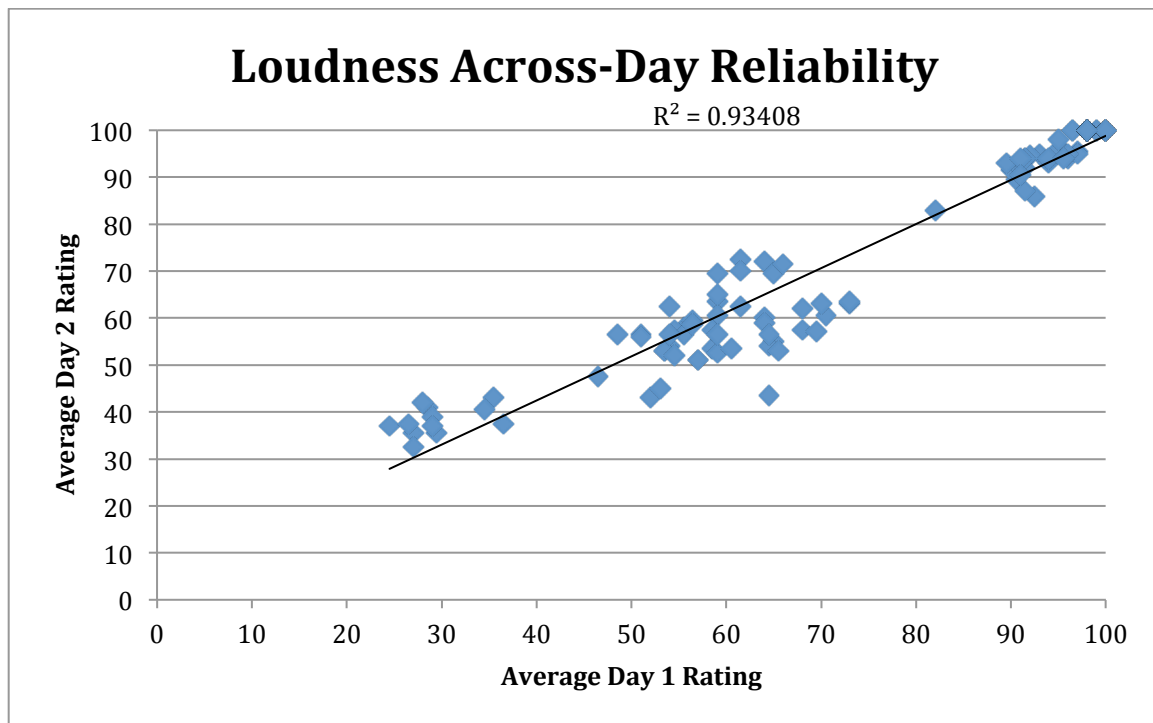


Fig. 7. Scatter plot of average adjustments on day one and the average adjustments on day two.

Fine-tuning adjustments across-day were somewhat reliable ($R^2 = 0.490$), but as seen in “within-day” reliability, listeners are less reliable with fine-tuning adjustments than with loudness adjustments across-day.

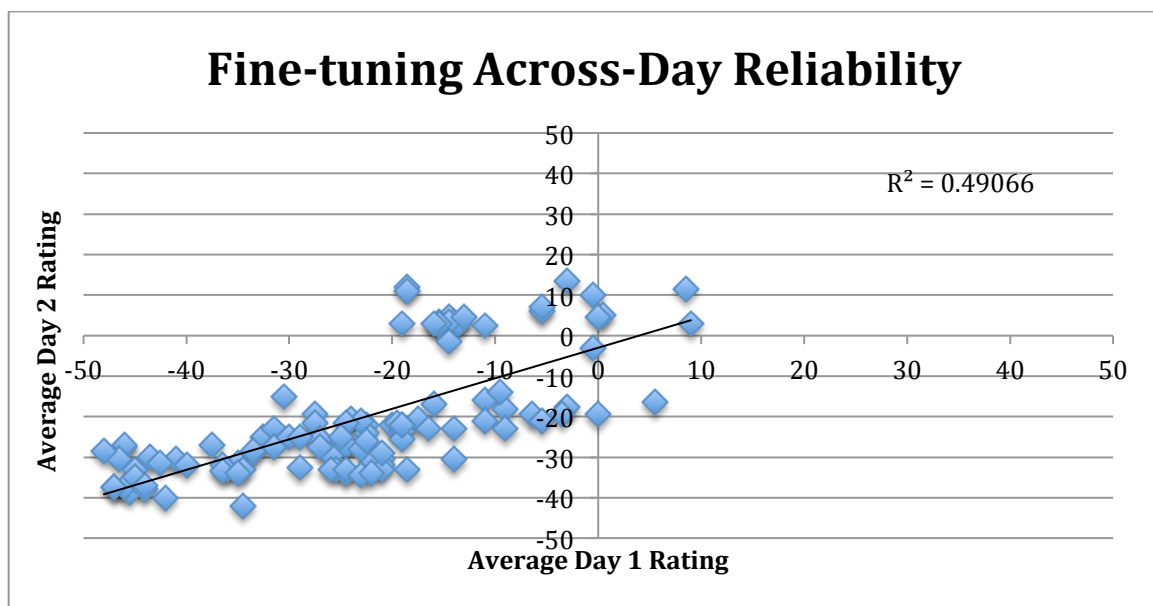


Figure 8. Scatter plot of average adjustments on day one and the average adjustments on day two.

While subjects were reliable within-day and across-day for making loudness adjustments, listeners were considerably less consistent with fine-tuning adjustments within-day and across-day.

Effects of genre, data-rate, and amount of DRC on user-adjustments

Figure 9 shows how listeners with normal hearing adjusted loudness settings for each of the twelve conditions. While the differences are small, individuals with normal hearing adjust the unprocessed recordings with less “loudness” than recordings with mild or heavy DRC, regardless of genre or data-rate. Subjects with normal hearing preferred to listen to both genres of music with mild DRC louder than music with heavy DRC.

On average, listeners with normal hearing preferred rock and classical MP3 files at 86 loudness; however, listeners with normal hearing prefer rock music slightly louder than classical music when at CD quality.

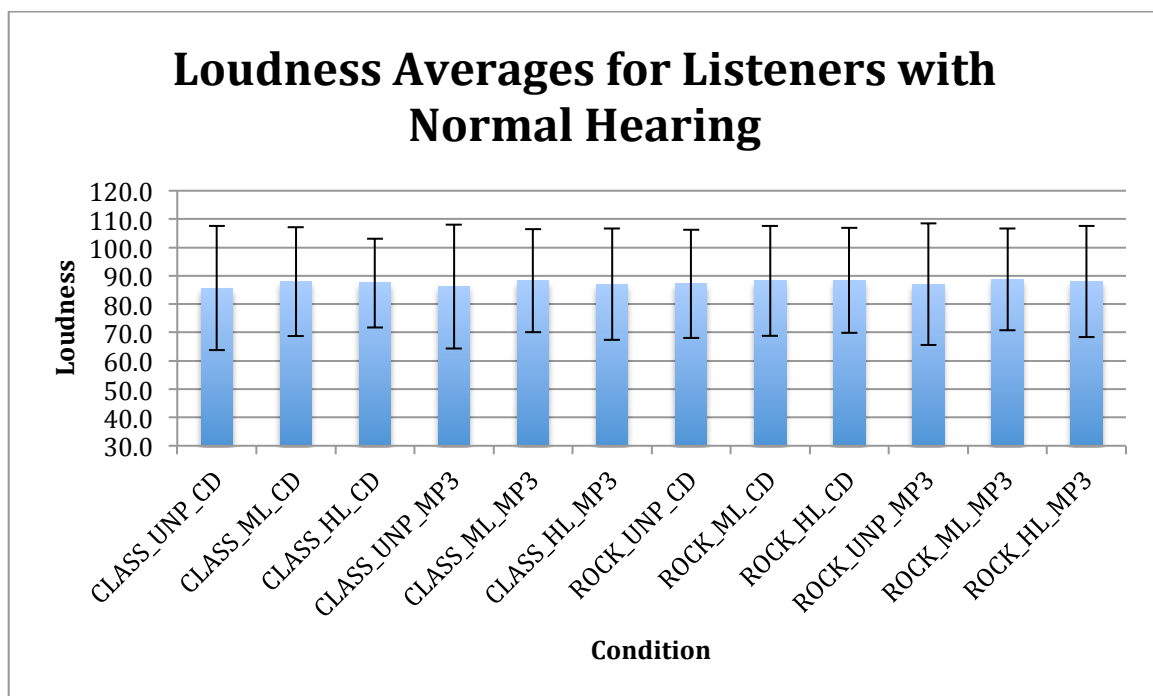


Figure 9. Average NH user-adjusted loudness settings across each 12 conditions.

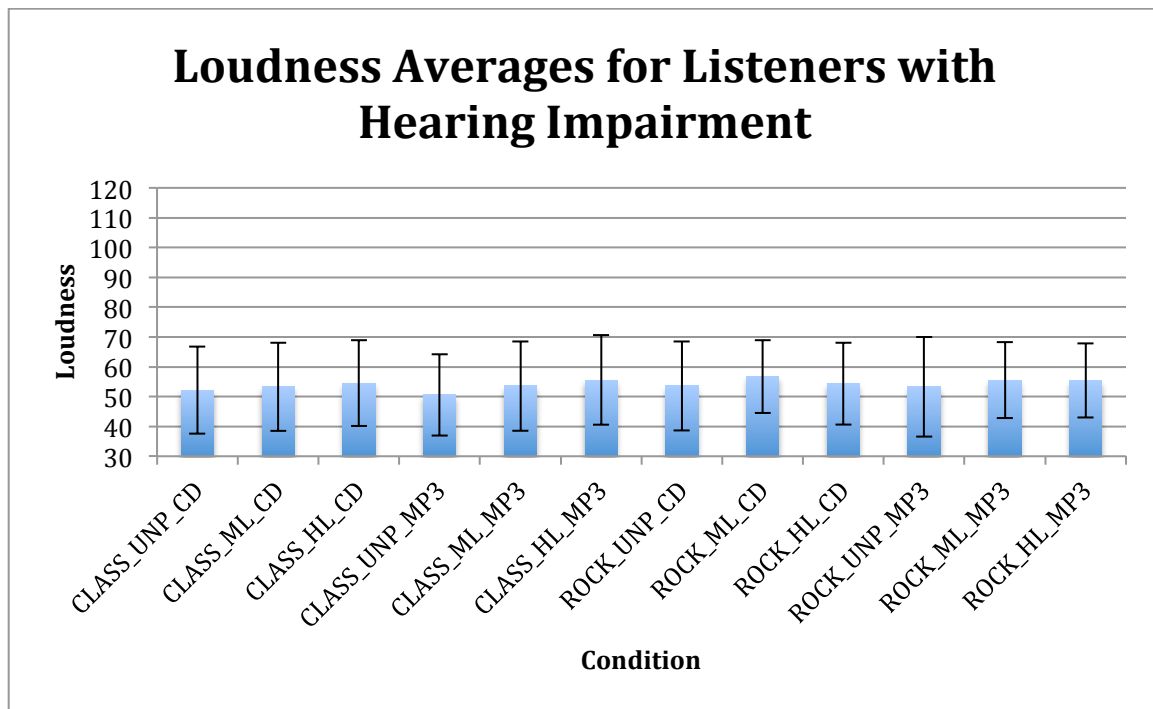


Figure 10. Average HI user-adjusted loudness settings across each 12 conditions.

On average, normal-hearing listeners prefer fine-tuning settings 4.5 units lower on MP3 files of rock compared to MP3 files classical. Normal hearing listeners prefer fine-tuning 2.5 units lower on CD quality files rock versus classical.

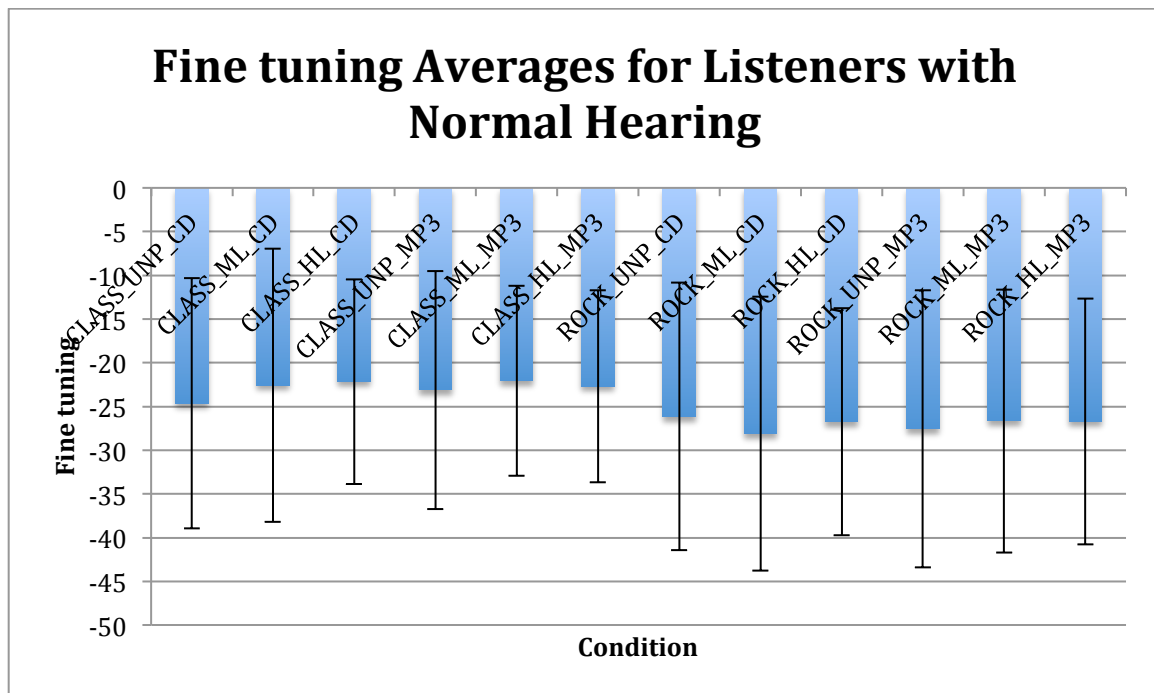


Figure 11. Average user-adjusted fine-tuning settings across each 12 conditions for listeners with normal hearing.

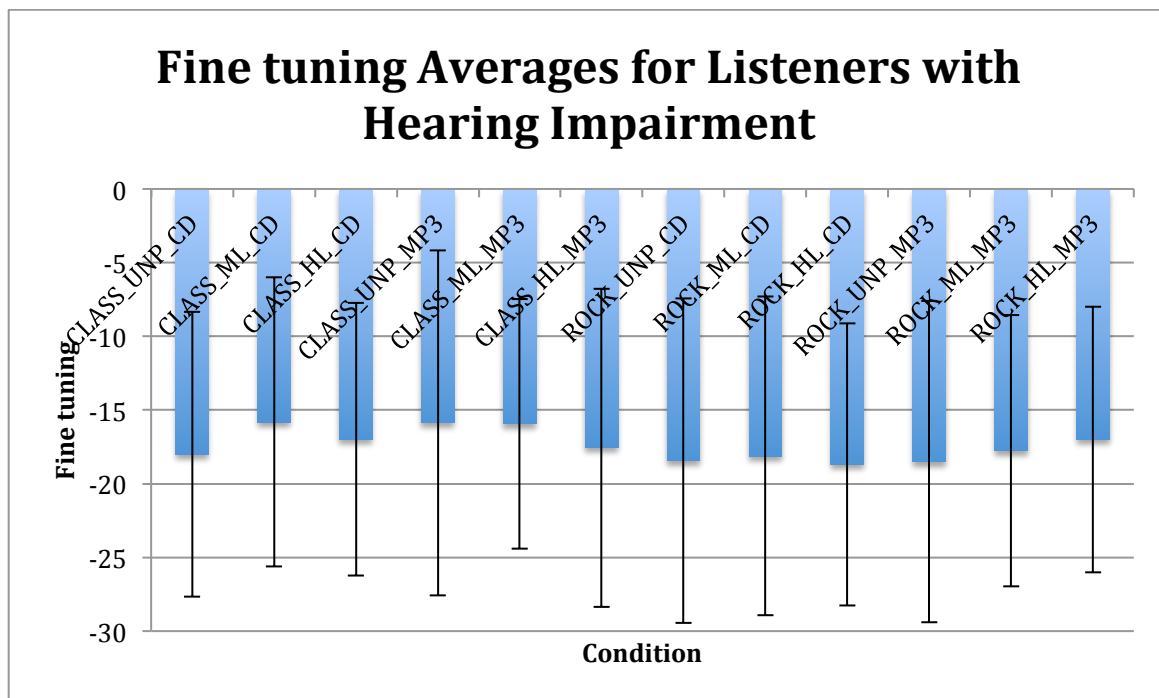


Figure 12. Average user-adjusted fine-tuning settings across each 12 conditions for listeners with hearing impairment.

Repeated Measures ANOVA

Factor	df	F	Sig.
Genre	1	2.623	.149
Data-rate	1	2.298	.173
DRC	2	.401	.677
Hearing-loss status	1	.681	.436

Table 6: Repeated measures ANOVA for loudness

We were interested in the effects of genre, data-rate, amount of DRC and hearing loss status on user-adjustments. For loudness adjustments, both amount of DRC and hearing loss status were significant; however, the power was quite low. Since the loudness controller applies WDRC to input

signals with heavy DRC, mild DRC and no DRC, it makes sense that the input signals' inherent amount of DRC affects the amount of WDRC (a second round of compression) that users chose to apply.

One reason why hearing-loss status is significant might be due to the differences in how WDRC (applied by the loudness controller) is designed to increase audibility for listeners with hearing impairment. Since the goal of WDRC,

Factor	df	F	Sig.
Genre	1	1.132	.323
Data-rate	1	.006	.943
DRC	2	5.109	.022
Hearing-loss status	1	8.369	.023

applied by the loudness controller, is to increase audibility for listeners with hearing impairment, there is no surprise that the hearing status of listeners is a significant effect on user-adjusted loudness.

As shown in Table 7, there were no significant factors in the way that listeners adjusted fine tuning. The current study's small number of subjects (9) could be a reason for this.

Discussion & Conclusions

The first research question addressed in this study was to assess whether listeners were reliable within and across days. The data presented here showed that subjects were reliable within-session in making loudness adjustments. Subjects also reliably adjusted loudness across sessions. These results are similar to the findings of other recent reports in the literature. For example, Sabin et al (2013) reported that participants are able to make reliable Ear Machine adjustments in about 10 seconds.

Fine-tuning adjustments were not nearly as reliable as loudness adjustments. The data presented here showed that subjects were only somewhat reliable in making fine-tuning adjustments within-session ($R^2 = 0.726$), and were even less reliable in making fine-tuning adjustments across-session ($R^2 = 0.490$). These results are different from the findings of other recent reports in the literature. Again, Sabin et al (2013) reported that participants are able to make reliable Ear Machine adjustments in about 10 seconds and did not make a distinction between loudness adjustments and fine-tuning adjustments.

One speculation for the difference in reliability between loudness adjustments and fine-tuning adjustments is that individuals have more background practice with adjusting music loudness on their car stereo, on their home stereo, and in other situations in life, whereas this experiment could very well be the first time a person will adjust spectral balance of music. Since people have had more practice in every-day life with adjusting loudness, their preference for loudness has been internalized more than their preference for spectral balance that an individual may have no previously acquired preferences for.

Individuals also made many more fine-tuning adjustments per condition than loudness adjustments for this reason, which is likely why the loudness adjustments were more reliable.

Another speculation is that the changes from one condition to another were brought out more by adjusting fine tuning than by adjusting loudness and thus, when subjects were presented a different condition, they felt the need to adjust fine tuning to be more imperative than loudness in adjusting the settings to get back to maximum sound quality. While both of these are just speculations, it is still unclear why the test-retest reliability for loudness adjustments within-day and across day was more reliable than the adjustments for fine tuning.

The second research question considered asked if listeners with normal hearing and listeners with hearing impairment adjust controllers differently for unprocessed, mildly compressed and heavily compressed music files. The results of the present study showed the different amounts of DRC to be a significant factor ($p < 0.05$) in the way that listeners adjust the loudness controller but not the fine-tuning controller. The presence of this significant effect is similar to the findings of Croghan et al (2012) that showed a significant effect of DRC on music listening preferences for rock and classical music because of the way that WDRC interacts with the DRC that is applied in the music industry. Croghan et al (2012) also demonstrated that heavy DRC significantly affected the perception of music. Since the loudness controller applies varying amounts of WDRC, there is no surprise that the findings of the current study correspond to the findings of Croghan et al (2012).

The third research question addressed in this study was to assess if listeners with normal hearing and listeners with hearing impairment adjust controllers differently for MP3 and CD quality files. The current study shows that there is a lack of significant effect for data-rate. This lack of significant effect is similar to the findings of other recent reports in the literature. For example, Pras et al (2009) found that listeners can hear a difference between CD quality and MP3 quality between 96-192 kbps, but only expert listeners can hear differences between CD quality and MP3 quality between 256-320 kbps. In the current study, we found that the factor of data-rate (128 kbps and 1411 kbps) did not affect listeners adjustments.

The fourth question addressed in this study was to assess whether listeners with normal hearing and listeners with hearing impairment adjust controllers differently for rock and classical music. The data presented here showed that genre did not have a significant effect on user-adjustments. This lack of significant effect contrasts with the findings of Croghan et al (2014). Croghan et al (2014) reported that the sound quality of classical music was decreased by heavy DRC, but for rock music, the amount of DRC was not a significant factor. The current study's findings are consistent with the findings of Higgins et al (2012), who reported that user preferences were the same for classical and rock music. Similarly, Moore et al (2011) found that user preferences were the same for classical and pop music. This leads one to believe that based on the current study's findings and past study's findings that genre was not a significant factor.

Lastly, the only between-subject factor, hearing loss status, was significant ($p < 0.05$) in how individuals adjusted loudness settings because individuals with hearing impairment are aided by WDRC differently than normal hearing listeners. Croghan et al (2014) found that subjects with hearing impairment preferred more linear WDRC. Three out of four of the subjects with hearing impairment preferred loudness between 50 and 70. The higher the loudness controller is set, more WDRC is applied and since WDRC also warps the signal envelope, reducing the spectral and temporal contrasts that listeners commonly utilize in music perception (Plomp, 1988), the current study's finding that listeners with hearing impairment prefer less WDRC coincides with the findings of Croghan et al (2014). The current study's findings are similar to the findings of Hansen et al (2002). Hansen et al (2002) reported that listeners with hearing loss perceived music signals differently than individuals with normal hearing.

While data-rate and genre were not significant factors in how users adjusted loudness, the amount of DRC and an individual's hearing status was a significant factor. This is largely due to the fact that the loudness controller applies compression in different amounts and when each condition has varying amounts of DRC, users tend to apply more or less WDRC (adjustments in the loudness controller) according to the amount of DRC that has already been applied to the input file in the music recording process. The conclusions of the current study are limited by not having specific compression ratio settings. The results of this study showed ratings; however, we did not look at user preference using a paired comparison, which future work might look into.

References

- Amlani, A. M. (2014). Apps for the EarsTablet and smartphone users can choose from a growing number of apps that test hearing sensitivity and amplify sound. *The ASHA Leader*, 19(7), 34-35.
- Amlani, A. M., Taylor, B., Levy, C., & Robbins, R. (2013). Utility of Smartphone-based Hearing Aid Applications as a Substitute to Traditional Hearing Aids. *The Hearing Review*, 20(13), 16-18, 20, 22.
- Arehart, K.H., Kates, J.M., & Anderson, M.C. (2011). Effects of noise, nonlinear processing, and linear filtering on perceived music quality. *International Journal of Audiology*, 50, 177-190.
- Atkinson, J. (2008). MP3 vs AAC vs FLAC vs CD. *Stereophile*, March, 31.
- Boymans, M., & Dreschler, W. A. (2011). Audiologist-driven versus patient-driven fine tuning of hearing instruments. *Trends in Amplification*, 1084713811424884.
- Bregitzer, L. (2009). *Secrets of Recording: Professional Tips, Tools & Techniques*. Focal Press.
- Brennan, R. L. (2010, August). Efficient embedded system algorithm mapping. In *Circuits and Systems (MWSCAS), 2010 53rd IEEE International Midwest Symposium on* (pp. 1065-1068). IEEE.
- Chasin, M., & Hockley, N. S. (2014). Some characteristics of amplified music through hearing aids. *Hearing research*, 308, 2-12.
- Chasin, M. (2003). Music and hearing aids. *Hearing Journal*, 56, 36–41.
- Chasin, M. (2006). Hearing aids for musicians. *Hearing Review*, 13, 11–16.
- Chasin, M. (2010). Amplification fit for music lovers. *Hearing Journal*, 63, 27–30.
- Chasin, M. (2010). Six ways to improve listening to music through hearing aids. *Hearing Journal*, 63(9), 27-30.
- Chasin, M. and Russo, F. (2004). Hearing aids and music. *Trends in Amplification*, 8, 35
- Christensen, K. T., & Jensen, R. G. (2012). *U.S. Patent Application 13/464,428*.
- Colucci, D. A. (2013). Hearing Matters: Aided Mapping for Music Lovers: Addressing the Basic Issues. *Hearing Journal*, 66(10).
- Cooper, D. H. (1982). Calculator program for head-related transfer function. *Journal of*

- the Audio Engineering Society*, 30(1/2), 34-38.
- Croghan, N. B., Arehart, K. H., & Kates, J. M. (2012). Quality and loudness judgments for music subjected to compression limiting. *The Journal of the Acoustical Society of America*, 132, 1177.
- Croghan, N. B., Arehart, K. H., & Kates, J. M. (2014). Music Preferences With Hearing Aids: Effects of Signal Properties, Compression Settings, and Listener Characteristics. *Ear and hearing*.
- Cröse, B., Kuk, F., & Bindeballe, H. O. (2011). Digital Wireless Hearing Aids, Part 4: Interference. *The Hearing Review*.
- Davies-Venn, E., Souza, P., & Fabry, D. (2007). Speech and music quality ratings for linear and nonlinear hearing aid circuitry. *Journal of the American Academy of Audiology*, 18(8), 688-699.
- Grimm, G., Guilmin, G., Poppen, F., Vlaming, M. S., & Hohmann, V. (2009). The personal hearing system: a software hearing aid for a personal communication system. *EURASIP Journal on Advances in Signal Processing*, 2009, 9.
- Hansen, M. (2002). Effects of multi-channel compression time constants on subjectively perceived sound quality and speech intelligibility. *Ear and Hearing*, 23(4), 369-380. doi: 10.1097/01.aud.0000028009.11739.3e
- Hawkins, D. B., & Naidoo, S. V. (1993). Comparison of sound quality and clarity with asymmetrical peak clipping and output limiting compression. *Journal of the American Academy of Audiology*, 4(4), 221-228.
- Higgins, P., Searchfield, G., & Coad, G. (2012). A comparison between the first-fit setting of two multichannel digital signal-processing strategies: Music quality ratings and speech-in-noise scores. *American Journal of Audiology*, 21, 13-21.
- Jenstad, L. M., & Souza, P. E. (2005). Quantifying the effect of compression hearing aid release time on speech acoustics and intelligibility. *Journal of Speech, Language and Hearing Research*, 48(3), 651.
- Jumisko, S., Ilvonen, V. P., & Väänänen-Vainio-Mattila, K. (2005, January). Effect of TV content in subjective assessment of video quality on mobile devices. In *Proceedings of SPIE* (Vol. 5684, pp. 243-254).
- Kates, J. M. (2008). Digital hearing aids. *Plural Publishing Inc*
- Kates, J. M. (2010). Understanding compression: Modeling the effects of dynamic-range compression in hearing aids. *International Journal of Audiology*, 49(6), 395-409.

- Katz, B., & Katz, R. A. (2003). *Mastering audio: the art and the science*. Butterworth-Heinemann.
- Keidser, G, O'Brien, A, Carter, L, McLelland, M, Yeend, I. (2008). Variation in preferred gain with experience for hearing-aid users. *International Journal of Audiology*, 47: 621-635.
- Kuk F, Lau C, Crose B, Keenan D, and Schumacher J. Efficacy of a Wireless TV Listening Device. *Hearing Review*. 2012; 19(09):38-45.
- Kuk, F., Crose, B., Korhonen, P., Kidmose, P., Jensen, M. H., Larsen, S. M., & Ungstrup, M. (2010). Digital Sampling. *Hearing Review*.
- Leek, M. R., Molis, M. R., Kubli, L. R., & Tufts, J. B. (2008). Enjoyment of music by elderly hearing-impaired listeners. *Journal of the American Academy of Audiology*, 19(6), 519-526.
- Markwalter, B. (2011). An Update from the Consumer Electronics Association (CEA). *SMPTE Motion Imaging Journal*, 120(6), 59-60.
- Moore, B. C. J., & Tan, C. T. (2003). Perceived naturalness of spectrally distorted speech and music. [Article]. *Journal of the Acoustical Society of America*, 114(1), 408-419. doi: 10.1121/1.1577552
- Moore, B. C., Glasberg, B. R., & Baer, T. (1997). A model for the prediction of thresholds, loudness, and partial loudness. *Journal of the Audio Engineering Society*, 45(4), 224-240.
- Moore, B. C. J., Fullgrabe, C., and Stone, M. A. (2011). Determination of preferred parameters for multichannel compression using individually fitted simulated hearing aids and paired comparisons. *Ear and Hearing*, 32, 556-568.
- Moore, B.C.J. (2012). Effects of bandwidth, compression speed, and gain at high frequencies on preferences for amplified music. *Trends in Amplification*, 16(3), 159-172.
- Moore, B. C. J., Tan, C. T., Zacharov, N., & Mattila, V. V. (2004). Measuring and predicting the perceived quality of music and speech subjected to combined linear and nonlinear distortion. *Journal of the Audio Engineering Society*, 52(12), 1228-1244.
- Moore, B. C. (1996, May). Masking in the human auditory system. In *Audio Engineering Society Conference: Collected Papers on Digital Audio Bit-Rate Reduction*.
- Ogihara, M., Takahashi, K., Taniguchi, S., & Mori, M. (2007, March). Networked Hearing Aids Using Wireless LAN Technologies. In *Wireless Communications*

- and Networking Conference, 2007. WCNC 2007. IEEE* (pp. 2680-2685). IEEE.
- Ovegard, A., Lundberg, G., Hagerman, B., Gabrielsson, A., Bengtsson, M., & Brandstrom, U. (1997). Sound quality judgement during acclimatization of hearing aid. *Scandinavian Audiology*, 26(1), 43-51.
- Parsa, V., Scollie, S., Glista, D., & Seelisch, A. (2013). Nonlinear Frequency Compression Effects on Sound Quality Ratings of Speech and Music. *Trends in Amplification*, 17(1), 54-68.
- Plomp, R. (1988). The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation-transfer function. *The Journal of the Acoustical Society of America*, 83, 2322.
- Pras, A., Zimmerman, R., Levitin, D., & Guastavino, C. (2009, October). Subjective evaluation of mp3 compression for different musical genres. *Audio Engineering Society Convention 127*.
- Puder, H. (2009, September). Hearing aids: An overview of the state-of-the-art, challenges, and future trends of an interesting audio signal processing application. In *Image and Signal Processing and Analysis, 2009. ISPA 2009. Proceedings of 6th International Symposium on* (pp. 1-6). IEEE.
- Ricketts, T. A., Dittberner, A. B., & Johson, E. E. (2008). High-frequency amplification and sound quality in listeners with normal through moderate hearing loss. *Journal of Speech Language and Hearing Research*, 51(1), 160-172. doi: 10.1044/1092-4388(2008/012)
- Rix, A. W., Beerends, J. G., Kim, D. S., Kroon, P., & Ghitza, O. (2006). Objective assessment of speech and audio quality—technology and applications. *Audio, Speech, and Language Processing, IEEE Transactions on*, 14(6), 1890-1901.
- Sabin, A. (2013, November 4). Why hearing app developers are forced to work with iOS instead of Android. RSS. Retrieved October 7, 2014, from <http://www.earmachine.com/blog/2013/11/4/why-hearing-app-developers-are-forced-to-work-with-ios-instead-of-android>
- Smartphone Control Apps for Hearing Devices: Caveat Emptor. (2013, November 19). *Hearing Economics Holly HosfordDunn Audiology Economics HearingHealthMattersorgHearingEconomics*. Retrieved October 7, 2014, from <http://hearinghealthmatters.org/hearingeconomics/2013/smartphone-control-apps-hearing-devices-caveat-emptor/>
- Souza, P. E., & Bishop, R. D. (2000). Improving audibility with nonlinear amplification for listeners with high-frequency loss. *Journal of the American Academy of*

Audiology, 11(4), 214-223.

Starch, D. (1908). Perimetry of the localization of sound. *The Psychological Review: Monograph Supplements*, 9(2), i.

Stone, M. A., Moore, B. C. J., Fullgrabe, C., and Hinton, A. C. (2009). "Multichannel fast-acting dynamic range compression hinders performance by young normal-hearing listeners in a two-talker separation task," *Journal of the Audio Engineering Society*. 57, 532–546.

Tan C.T. & Moore B.C.J. 2008. Perception of nonlinear distortion by hearing-impaired people. *International Journal of Audiology*, 47, 246 – 256.

Tan, C. T., & Moore, B. C. J. (2008). Perception of nonlinear distortion by hearing-impaired people. *International Journal of Audiology*, 47(5), 246-256. doi: 10.1080/14992020801945493

Tan C.T., Moore B.C.J., Zacharov N. & Mattila V.V. 2004. Predicting the perceived quality of nonlinearly distorted music and speech signals. *Journal of the Audio Engineering Society*, 52, 699 – 711.

Van Buuren R.A., Festen J.M. & Houtgast T. 1999. Compression and expansion of the temporal envelope: Evaluation of speech intelligibility and sound quality. *The Journal of the Acoustical Society of America*, 105, 2903 – 2913.

Villchur, E. (1973). Signal processing to improve speech intelligibility in perceptive deafness. *The Journal of the Acoustical Society of America*, 53, 1646.

Vickers, E. (2011). The loudness war: Do louder, hypercompressed recordings sell better?. *Journal of the Audio Engineering Society*, 59(5), 346-351.

What Does A Song That Costs \$5 Sound Like? : The Record : NPR. (n.d.). *NPR : National Public Radio : News & Analysis, World, US, Music & Arts : NPR*. Retrieved September 24, 2013, from <http://www.npr.org/blogs/therecord/2013/09/11/219727031/what-does-a-song-that-costs-5-sound-like>